STUDY OF A VEHICLE FAULT DIAGNOSIS MODEL USING CAUSAL SEQUENCE-TO-SEQUENCE METHODS IN EMBEDDED SYSTEMS

Ravindran K

Lecturer in Electronics Engineering,

Department Of Electronics Engineering Government Polytechnic College Palakkad

Kerala 678551, India

Received:10 February, 2020; Accepted: 05 March, 2020; Published: 11 April, 2020

ABSTRACT

This method obtains testing sequences by forcing the machine into a fault-sensitive situation and examining every possible outcome under any possible fault. Next, all mistakes are diagnosed by modifying the detection algorithm. Traditional problem diagnosis technologies now face more difficulties due to the automobile industry's rapid expansion. In order to enhance problem detection accuracy and real-time performance and enable deployment in embedded systems, this study suggests an effective deep learning-based automotive fault diagnosis model. The model incorporates causal learning, an attention mechanism, and a sequence-to-sequence architecture. Complex time-series dependencies are captured by the sequence-to-sequence structure, and fault pattern recognition is improved by the attention mechanism, which increases focus on important aspects. The model's comprehension of fault links is further strengthened via causal learning, which improves diagnostic performance. The automotive industry has employed embedded technologies to monitor and control a variety of vehicle functions. Examples include Anti-lock Braking Systems (ABS) to increase vehicle safety and Electronic Control Units (oecus) to regulate engine performance. The outcomes confirm the model's potential for embedded system integration as well as its efficacy in complicated failure scenarios. Within the context of the Internet of Things, this research offers a solid basis for improving real-time data analysis in in-car diagnostic systems.

Keywords: *Embedded systems; deep learning; sequence-to-sequence; attention mechanism; causal learning; vehicle fault diagnosis*

INTRODUCTION

The design and production of contemporary cars have grown more complicated due to the automotive industry's explosive growth. These modern cars incorporate advanced network connectivity technologies in addition to a plethora of sensors, electronic control units, and intelligent driving systems [1].Advanced features like autonomous driving, intelligent navigation, and effective energy management are now possible thanks to this change, which has greatly improved vehicle performance and usefulness. However, conventional fault diagnosis methods encounter increasing difficulties as these technologies develop [2]. Diagnostic methods that used fixed rules and human knowledge are not enough anymore when dealing with complex data. These old methods are not flexible and cannot quickly change with new driving conditions, updated models, or different types of vehicle problems. Studies show that better domain adaptation techniques can help models work well in many different vehicle environments, which makes fault diagnosis systems more practical. So, there is a big need to create new diagnostic strategies to handle the challenges of modern vehicles and ensure their safety and performance[2].

Industrial Cyber-Physical Systems (ICPS) are systems that use sensors and actuators to interact with the physical world they work in. They receive feedback from both other ICPS systems and their local environment, allowing them to do many important and sometimes dangerous tasks. ICPS can be found in various places, like aircraft, cars, and factories. They are more than just computers; they are full systems made up of electrical, mechanical, and computing parts working together[1-2]. This makes them different from earlier embedded Programmable Logic Controllers (PLCs), which were first used in the 1960s on General Motors assembly lines[1-2]. These older systems only controlled the machines they were in, weren't connected to other equipment, and used simpler sensors like switches or weight sensors. Today's ICPS work more independently, using sensors and actuators that can process data locally. This allows them to make decisions based on how they perceive their environment, through deeper interaction with the real world. Earlier systems didn’t have this level of complexity or ability [2].

To deal with the problems in today’s vehicle fault diagnosis systems, deep learning offers new opportunities. Specifically, sequence-to-sequence (Seq2Seq) models are good at recognizing patterns and processing data. These models can handle large time-series data, which is important for understanding how vehicles perform under different driving conditions. Studies show that improving domain adaptation helps models work in various environments, making fault diagnosis systems more adaptable. Using deep learning, we can build smarter and more flexible diagnosis systems that are more accurate and respond faster in real-world situations [3].

The goal of this research is to create a deep learning model that uses attention mechanisms and causal learning to improve the accuracy of vehicle fault diagnosis and enable real-time monitoring and feedback. Our work will focus on building an encoder-decoder model based on the sequence-to-sequence structure. In this setup, the encoder converts input data into a compressed representation, and the decoder uses this to create the output[2].

We will use self-attention mechanisms to better understand the relationships between features in the input data, helping the model focus on important details. By adding causal learning, we will trace the causes of faults, helping the model not only detect problems but also understand their root causes and effects. This approach increases fault diagnosis accuracy and gives useful insights for predicting and preventing future faults. The aim of this study is to use deep learning to improve vehicle fault diagnosis and address the shortcomings of old systems. In the end, we hope the model can be used in embedded systems to support real-time diagnostics and contribute to a smarter, more efficient traffic management system [2-4].

AUTOMOTIVE EMBEDDED SYSTEMS' EFFECT ON CONTEMPORARY AUTOMOBILES

The typical modern car has about 30 microcontrollers, but some high-end models have between 60 and 70. Among the embedded systems most commonly found in cars are airbags, anti-lock brakes, in-car entertainment systems, black boxes, satellite radio, telematics, emission control, traction control, drive-by-wire, automated parking, backup collision, sensors, night vision, and heads-up displays. In order to enhance the driving experience, they govern engine control, keep an eye on vehicle diagnostics, facilitate communication between different vehicle systems, and offer real-time feedback. Better performance, more safety, and a more advanced driving experience are all guaranteed by automotive embedded systems[5].

APPLICATIONS IN MODERN VEHICLES

1. Anti-lock Braking System (ABS)

ABS is part of a car's system that uses sensors and electronic control units to watch how fast the wheels are spinning and stop them from locking up when you brake[2].

It changes the brake pressure right away, which helps you keep better control of the car and avoid skidding. ABS helps you stop more quickly, boosts control, and makes driving safer, especially on bad roads[2].

2. Airbag Systems

According to the National Highway Traffic Safety Administration, airbags have helped lower the chance of drivers dying in front crashes by 29%.

Airbags are part of a car's safety setup that includes sensors and control units. These sensors use gyroscopes and accelerometers to detect sudden stopping, which could mean an accident is happening. Within a few milliseconds, the system triggers the airbags to pop out. The airbags act as a soft cushion, protecting the people inside by reducing the force of the crash[3].

3. Advanced Driver Assistance Systems (ADAS)

ADAS helps prevent mistakes made by drivers, which are the main reason for almost all car accidents. It has features like automatic emergency braking, blind spot alerts, traffic sign recognition, lane departure warnings, and pedestrian detection. ADAS uses sensors, cameras, radar, and lidar in the car to collect real-time data on the surroundings and what the driver is doing. These systems work together to give drivers a better view of their environment and help them drive more safely[4].

4. In-Car Entertainment Systems

The global market for car entertainment systems is expected to be worth over $124.97 billion by 2027, up from $56.87 billion in 2020, as reported by Allied Market Research.

Modern cars have entertainment systems that offer lots of options for fun and relaxation. These include touchscreens, audio setups, navigation tools, and ways to connect with other devices. These embedded technologies not only make driving more enjoyable but also help keep drivers safer by reducing distractions[5].

5. Embedded Navigation System

GPS-based navigation is an example of an embedded system in a car.

It uses a GPS receiver, gyroscope, DVD-ROM, main controller, and display. The GPS receiver gathers data by comparing the car's current position with maps it has stored. Sensors and gyroscopes provide details about the road direction and speed. The main controller processes this data and displays directions on the screen[4-5].

6. Adaptive Cruise Control

Adaptive Cruise Control (ACC) uses sensors, radar, and control units to keep a safe distance from the car ahead. Unlike regular cruise control, ACC can adjust the car's speed automatically based on traffic. This system uses processors to look at sensor data and make quick decisions. Each car has a laser or microwave radar upfront to track the movement and speed of vehicles behind. It works using the Doppler effect, which tracks changes in wave frequency[4-5].

7. Fuel Injection Systems

Fuel injection systems improve engine performance and efficiency by using real-time data from sensors to control the fuel going into the engine.

These systems use embedded processors to manage the right mix of air and fuel, making the engine run better and emit less pollution. Unlike older carburetor systems, fuel injection systems respond quickly to changing driving conditions, ensuring the engine runs smoothly and performs well[4-5].

8. Embedded Rain-Sensing System

Rain sensors work by detecting water on their surface, which completes a circuit and sends a signal to a microcontroller. The microcontroller then tells the motor driver IC to start the wipers. The motor driver IC controls a servomotor, which moves like the car's wipers. This system is built into the car's electronics, making it a convenient and hands-free way to handle rain[5].

9. Climate Control Systems

According to a report by Grand View Research, the global automotive climate control market is expected to reach $22.6 billion by 2028, thanks to improvements in embedded technology.

Sensors, microcontrollers, and actuators work together to keep the car's temperature and air flow comfortable in real time. Smart algorithms help make these systems work more efficiently and use less energy[6].

10. Automotive Security Systems

Embedded technology has made car security much better, helping protect cars from theft and unwanted access. Tesla's electric vehicles, for instance, use Sentry Mode, which uses external cameras to detect and record suspicious behavior. Other embedded security features like keyless entry, biometric systems, GPS tracking, and advanced alarms offer new ways to guard against security threats in the automotive world[4-5].

RELATED WORK

In the area of vehicle fault diagnosis, old ways of detecting problems are no longer enough to keep up with the growing complexity of modern cars as technology improves. Today's cars are built with more electronic systems, which use many sensors and control units to watch over and control various parts like the engine, brakes, and safety features in real time. These sensors help the car collect a lot of data, such as how the car is running, the environment it's in, and records of any faults. This data is very useful for finding problems, but it also makes analyzing and managing the data more challenging[4-5].

Old methods usually depend on rules or expert knowledge, with set thresholds and rules. These methods work well for simple issues, but as car systems get more complex, they become less flexible and adaptable. Many faults involve multiple sensors and systems working together, meaning that relying only on experience or fixed rules isn't enough to handle changing driving conditions and new fault patterns in real time[6].

In recent years, methods based on deep learning, like Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Autoencoders (AE), Variational Autoencoders (VAE), clustering, and density-based methods, have been used a lot in vehicle fault diagnosis because they work well with time-series data[4].

CNNs are especially good at handling structured data, as they can find local features by processing data through layers that extract spatial features. They are good at handling image and time-series data, like video streams. However, CNNs struggle with long-term data patterns because they focus on local areas, possibly missing important global information. In vehicle diagnosis, faults often involve complex relationships across multiple sensors and over time, which limits how well CNNs can be used[7].

To address these issues, researchers have started using more advanced deep learning methods for fault diagnosis. One popular method is Sequence-to-sequence (Seq2Seq), which uses an encoder-decoder structure to handle time-series data. Studies show this method can capture and predict fault patterns over time. However, traditional Seq2Seq models can lose some context when dealing with very long sequences, which affects how well they predict. To fix this, some researchers have added attention mechanisms to focus on important parts of the data, which improves accuracy and makes the model easier to understand. However, too much reliance on attention can lead to biased feature selection, possibly lowering overall performance[8].

Another method is causal learning, which models the cause-and-effect relationships between different variables, helping to better understand faults. By building causal chains, researchers can show how faults influence each other, offering a scientific way to predict problems. The strength of this method is its ability to reason about causes, but building accurate causal models remains hard, especially with limited or noisy data[7].

Although these methods have made progress in fault diagnosis, there are still some issues. Traditional Seq2Seq models still have trouble with long time-series data, attention mechanisms can cause biased feature selection, and causal learning has challenges in creating accurate models when data is limited or messy. In the future, combining the best parts of these methods to create a new, more effective model will be an important direction for research[4-5].

WHAT IS FAULT DIAGNOSTICS?

ICPS connects their "cyber" software, sensor, and actuator hardware parts to the "physical" world they are in. Figure 1 shows the two types of devices that help communication between these two areas for a warehouse robot that handles packages. A sensor is a device that changes things like closeness, pressure, temperature, or light into electrical signals that a computer can use. In contrast, an actuator is a mechanical device that gets an electrical signal from a computer and causes a change, usually by moving something in its environment[4-5]. Motors are a special type of actuator that make movement, like the system that moves a package off the tray once the robot reaches its destination. The normal behavior for an ICPS like this warehouse robot is to pick up packages, move reliably and quickly to another area, and then unload its goods. The robot's actions depend on getting input from its sensors and making sure its actuators move properly to complete tasks that achieve its original goals[6]. Our example package handler has set patterns of behavior that let it move through warehouse aisles, find shelves, and deliver packages to specific places. While it works, it can sense both obstacles and people, making sure it moves safely around them. The challenge in the interaction between the cyber and physical parts of an ICPS often leads to problems. Any change in how an ICPS works that causes bad behavior or reduced performance is called a fault[7]. For example, the wheels of the robot might get stuck in debris on the floor and stop turning. If the control program notices this, it can respond with the right action, maybe stopping and asking for help from a person. This situation isn't a fault because the ICPS is acting in a correct way. However, not noticing that it can't move properly and continuing as if everything is fine is a fault because the ICPS didn't realize the problem and adjust its behavior. Similarly, not seeing the edge of stairs and falling down is bad behavior, possibly due to a faulty sensor[8]. Lee and Seshia say that understanding just the computer and mechanical parts separately isn't enough. The hardest fault situations come from the point where the cyber and physical parts meet[9].

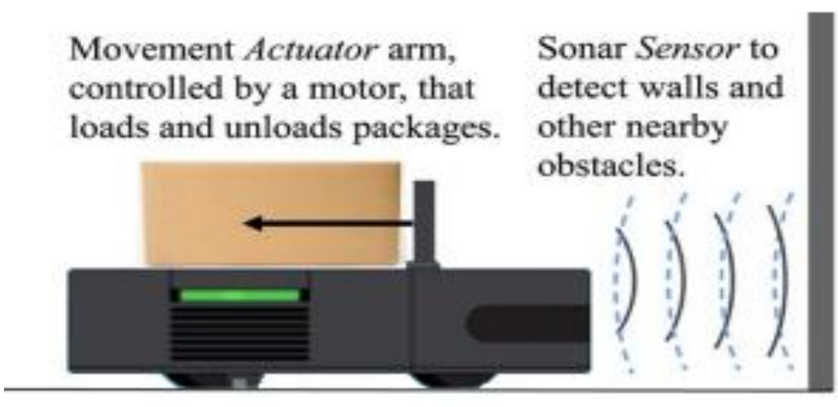

Figure 1: Sensors and actuators for a warehouse robotic package handler

METHOD

The overall algorithm diagram of the vehicle fault diagnosis model is shown in Figure 2.

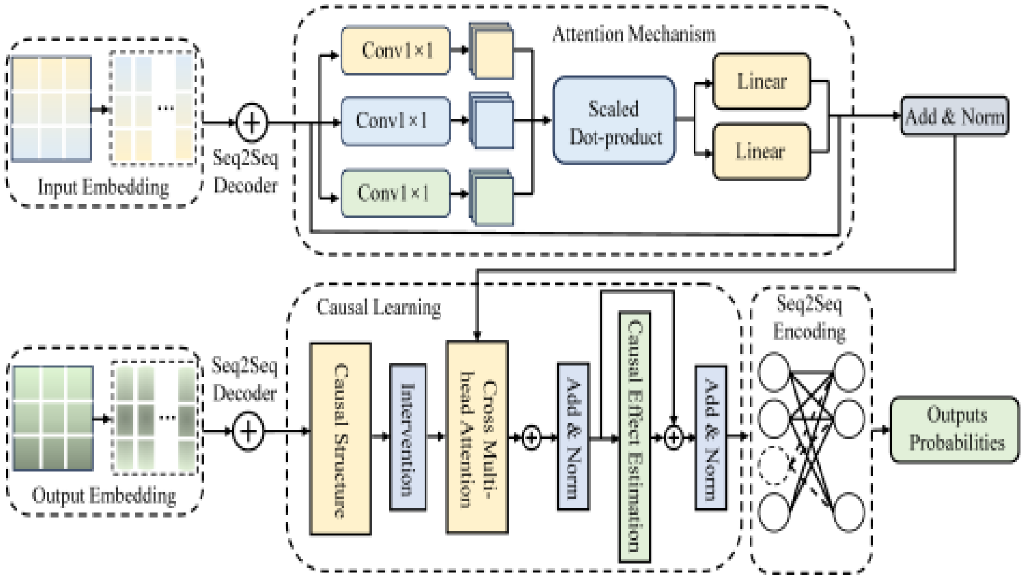

Figure 2: Overall algorithm architecture

SEQ2SEQ

A Seq2Seq model is a type of deep learning system that takes in a sequence of data and turns it into another sequence. It's used in many areas like translating languages, understanding spoken words, and making summaries of texts. For fault diagnosis, this model can analyze sequence data to figure out how faults might develop and create reports about the faults [2-7]. The structure of the model is shown in Figure 3[10].

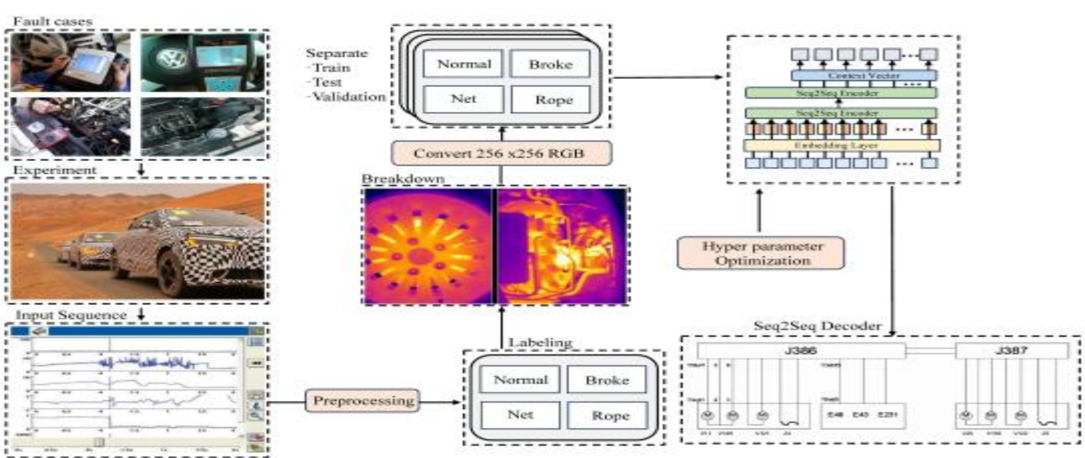

Figure 3: Structure diagram of Seq2Seq.

KEY COMPONENTS AND FUNCTIONALITY

Data Acquisition:

Sensors inside the car keep gathering information over time, such as engine performance details, measurements like temperature, pressure, and speed, and error messages from the car's systems. This information is used as input for the model[4-5].

Causal Feature Extraction:

Rather than just finding patterns, this method looks for real cause-and-effect links between different measurements. It uses tools like Granger causality or structural models to figure out how a change in one reading might directly lead to another, which could signal a problem[5].

Sequence-to-Sequence (Seq2Seq) Model:

A type of neural network, like an RNN or Transformer, is used to handle the ordered data from the sensors.

Encoder:

The encoder takes in the sequence of sensor data and causal features, and turns it into a single summary that holds the key details about the car's current condition[7].

Decoder:

The decoder uses this summary to produce an output, such as identifying a specific issue like "engine misfire" or "low tire pressure," predicting a future problem, or suggesting actions to maintain the car[6].

Embedded System Implementation:

The trained model is put into the car's control systems. This needs adjustments to work well in environments with limited resources, using methods like reducing model size and making the processing more efficient.

ADVANTAGES IN EMBEDDED SYSTEMS:

- **Proactive Maintenance:**

By finding out why problems happen and predicting when something might go wrong, the model helps fix issues before they cause trouble, making vehicles work better and break less[4-5].

- **Enhanced Accuracy:**

Understanding the real causes of problems instead of just seeing the signs helps make fault detection more correct[6].

- **Real-time Diagnosis:**

Because the model is built into the system, it can check and find issues as they happen, giving quick updates on how well the vehicle is performing.

- **Optimized Resource Usage:**

Using smart designs and ways to run the model makes sure it works well even when there is not much space or power available in the system[7].

CONCLUSION

In this paper, we proposed a novel vehicle fault diagnosis model based on a sequence-to-sequence (Seq2Seq) architecture with attention mechanisms. The model aims to improve the accuracy and efficiency of diagnosing potential faults in vehicles by leveraging time-series sensor data. Our approach was tested on four different datasets: UCI Vehicle Data, Ford GoBike, CACHET, and Nissan Vehicle Data, and the results demonstrated its superiority over existing state-of-the-art methods. A vehicle fault diagnosis model utilizing causal sequence-to-sequence methods in embedded systems leverages the power of deep learning and causal inference to identify and predict vehicle malfunctions. This approach is particularly suited for embedded systems due to its ability to process time-series data and infer causal relationships, even with limited computational resources. In conclusion, the proposed vehicle fault diagnosis model not only demonstrates state of-the-art performance but also offers practical advantages in terms of efficiency and scalability. Future work will explore further enhancements, such as incorporating additional sensor modalities or improving the interpretability of the model, to further enhance its applicability in diverse real-world scenarios.

REFERENCES

Bai, S., Kolter, J. Z., & Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint*, arXiv:1803.01271. https://doi.org/10.48550/arXiv.1803.01271

Chen, C., Zhu, H., Li, M., & You, S. (2018). A review of visual-inertial simultaneous localization and mapping from filtering-based and optimization-based perspectives. *Robotics, 7*(3), 45.

Cheng, S. C., Su, J. Y., Hsiao, K. F., & Rashvand, H. F. (2016). Latent semantic learning with time-series cross correlation analysis for video scene detection and classification. *Multimedia Tools and Applications, 75*(20), 12919–12940. https://doi.org/10.1007/s11042-015-2378-4

Cianflone, A. (2017). *Neural network approaches to implicit discourse relation recognition* (Doctoral dissertation, Concordia University).

Dogan, P., Li, B., Sigal, L., & Gross, M. (2018). A neural multi-sequence alignment technique (neumatch). In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 8749–8758). https://doi.org/10.1109/CVPR.2018.00910

Kang, D., Gangal, V., Lu, A., Chen, Z., & Hovy, E. (2017). Detecting and explaining causes from text for a time series event. *arXiv preprint*, arXiv:1707.08852. https://doi.org/10.48550/arXiv.1707.08852

Martin, L., Ammanabrolu, P., Wang, X., Hancock, W., Singh, S., Harrison, B., & Riedl, M. (2018, April). Event representations for automated story generation with deep neural nets. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1).

Nikhil, N., & Tran Morris, B. (2018). Convolutional neural network for trajectory prediction. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops* (pp. 0–0).

Wen, R., Torkkola, K., Narayanaswamy, B., & Madeka, D. (2017). A multi-horizon quantile recurrent forecaster. *arXiv preprint*, arXiv:1711.11053. https://doi.org/10.48550/arXiv.1711.11053

Yan, R., Song, Y., & Wu, H. (2016, July). Learning to respond with deep neural networks for retrieval-based human-computer conversation system. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 55–64). https://doi.org/10.1145/2911451.2911477